IDA

INSTITUTE FOR DEFENSE ANALYSES

Assessment of SIMITAR: Status Report Two (NTC98)

John Metzko
John Morrison

October 1998

IDA Document D-2217
Log : H 98-002935

This work was conducted under contract DASW01 94 C 0054, Task A-132, for the Defense Advanced Research Projects Agency. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

PREFACE

This report is the second in partial fulfillment of a task entitled "Advanced Distributed Simulation Technology," which is sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Army National Guard (ARNG). It describes the effectiveness of Simulation in Training for Advanced Readiness (SIMITAR) in the performance of the 116th Armored Cavalry Brigade, one of two SIMITAR test brigades, at the National Training Center (NTC) in July 1998. A previous report described the effectiveness of the other SIMITAR test brigade, the 48th Mechanized Infantry, at the NTC 2 years earlier.

The plan for evaluating SIMITAR is focused on measuring training results. Since SIMITAR trains numerous tasks, we want to measure performance of those tasks in the field. The critical part of the evaluation involves grading task performance, which can be done credibly only by Army observer controllers (OCs). Thus, we gratefully thank the OCs of several NTC teams—Broncos, Cobras, Goldminers, Scorpions, Sidewinders, Tarantulas, and Werewolves—for their essential help in evaluating SIMITAR at NTC98.

We thank IDA reviewers Dexter Fletcher and Michael Rigdon, whose suggestions prompted us to add material that, we hope, will make the report more comprehensible to readers outside the ARNG-Army training cognoscenti. And we also thank the following external reviewers for their comments and suggestions: Peter Grundvig, SIMITAR Project Officer with the 116th Cavalry Brigade; Major Al Fracker, U.S. Army Infantry School (formerly SIMITAR Project Officer with the 48th Mechanized Infantry Brigade); and LTC Tom Harrington, Commander of the 1-163 Mechanized Infantry Battalion, Montana Army National Guard.

CONTENTS

TABLES

EXECUTIVE SUMMARY

This evaluation of the effectiveness of SIMITAR (Simulation in Training for Advanced Readiness) at NTC98 (National Training Center in 1998) involves primarily two of four training areas in the SIMITAR Program: BSS (Battle Staff Synchronization) and CSS (Combat Service Support). While the other two training areas, Gunnery and Maneuver, will be covered in a subsequent overall SIMITAR assessment report, we are able to make a limited assessment of Platoon Maneuver at NTC98.

A. INTERVENTIONS, TASKS, AND BASELINES

SIMITAR uses various "interventions"—simulation devices, computer-based courseware, and training strategies—to train tasks, which are drawn from ARTEP (Army Training and Education Program) MTPs (Mission Training Plans). Performance of tasks, over 70 percent of which are at the brigade and battalion levels, is rated by NTC OCs (observer-controllers) using the Army's 3-point grading system: T (Trained), P (Need Practice), and U (Untrained). The task list includes 403 brigade- and battalion-level tasks, 68 company-level tasks; 47 platoon-level tasks, most of which (45) applied to Scout and Mortar platoons; and 117 tasks for tank and BFV (Bradley Fighting Vehicle) platoons.

Task performance by 116th Armored Cavalry Brigade units at the NTC in July 1998 was compared to a baseline performance by 48th Mechanized Infantry Brigade units at the NTC in June 1996. The baseline tasks were similar except for platoon tasks, the performance of which was not evaluated at NTC96. While both are SIMITAR test brigades, the 116th Brigade had a 10-to-1 advantage in SIMITAR training hours over the 48th Brigade in 2–3 years of training before their NTC rotations. While the

NTC98 and NTC96 performance comparison is not a well-controlled experiment, we know of no factor that invalidates the results (see Chapter I.C).

The above comparison uses an external baseline, i.e., the performance of other Army National Guard (ARNG) units training at the NTC, to benchmark performance of the 116th Brigade at NTC98. That some tasks are trained by SIMITAR devices and courseware while other tasks are not gives us another benchmarking option. We use this within-unit baseline to assess SIMITAR effectiveness at the brigade, task force/battalion, and company levels for performance at both NTC98 and NTC96. And since platoon training performance was not measured at NTC96, we use the internal baseline method to evaluate SIMITAR interventions associated with platoon Maneuver performance at NTC98.

B. COMBAT ELEMENTS

Performance at both NTC98 and NTC96 was best at the platoon level and decreased progressively at higher echelons—company, battalion, and brigade. NTC98 performance was much better at all echelons than performance at NTC96 (see Tables III-4 and III-5). Tank and BFV platoons sustained their performance at AT97 (Annual Training 1997) in the more challenging NTC environment (Table III-8). In every case in which within-unit baselines are used, average performance on SIMITAR-trained tasks was about equal to or significantly better than performance on tasks not trained by SIMITAR (Tables III-3 and III-7).

1. Brigade

Brigade performance at NTC98 was significantly better in four of six Battlefield Operating Systems (BOSs)—Intelligence, Air Defense, Mobility and Survivability, and CSS—than Brigade performance at

NTC96 (Table III-9).[1] By statistical significance tests, performance in the other BOSs at NTC98 and NTC96 were not significantly different.

Of the SIMITAR interventions that can train brigade-level tasks (some same and some different) in the various BOSs only one, Janus, was actually used for Brigade training. The Brigade at NTC98 had a 12-to-1 advantage in the number of Janus exercises over the NTC96 Brigade in pre-NTC training.

2. Armored Task Force

Armored TF (Task Force) performance at NTC98 was also significantly better in four of seven BOSs—Maneuver, Fire Support, C2 (Command and Control), and CSS—than Armored TF performance at NTC96 (Table III-10). Performance in the other BOSs at NTC98 and NTC96 were not significantly different by statistical significance tests.

Here also, of the seven SIMITAR interventions that can be used to train tasks (some same and some different), only one, Janus, was actually used for Armored TF training (illustrated by Table III-12). The Armored TF at NTC98 had a 12-to-3 advantage in the number of Janus exercises over its NTC96 counterpart in pre-NTC training. Thus, we see evidence that Janus was responsible for the superior training performance of the Brigade and the Armored TF.

3. Mechanized Infantry Task Force

Performance results for the Mech Infantry TF at NTC98 were much different from those for the Brigade and Armored TF. The NTC98 performance of the Mech Infantry TF was significantly better in only two BOSs—Fire Support and Air Defense—than that of the NTC96 Mech Infantry TF. Moreover,

---

1 BOSs are the major battlefield functions performed by the Army to execute its operations. There are seven: (1) Maneuver, (2) Fire Support, (3) Air Defense, (4) Command and Control (C2), (5) Intelligence, (6) Mobility and Survivability, and (7) Combat Service Support. While all seven BOSs apply to task force operations, only the last six apply to the brigade, which does not maneuver at the NTC.

the latter TF was significantly better in four BOSs—Maneuver, C2, Mobility and Survivability, and CSS—than the Mech Infantry TF at NTC98 (Table III-11). Since Janus training was the dominant factor in Brigade and Armored TF performance, we look there for a possible explanation of the Mech Infantry TF performance.

The 116th Brigade's Mech Infantry TF, whose core is the 1-163 Mech Infantry Battalion, was formed in FY95 by combining several disparate Montana ARNG units (Chapter III.A.5). The 1-163 Battalion spent FY96 on NET (new equipment training). That left 57 training days in the less than 2 years available to train for the NTC rotation in July 1998. In its pre-NTC training the 1-163 Battalion staff, which was the Mech Infantry TF staff, had eight TF-level Janus exercises compared to four for the 48th Brigade's Mech Infantry TF at NTC96. However, the nominal Janus training advantage was nullified by Battalion/TF staff changes—S1, S2, S3, Assistant S3, S4, and Fire Support Officer—in the year between AT97 and NTC98. Among them, these six new staff members missed two-thirds of the BSS training in the eight Janus exercises. The counterpart Mech Infantry TF staff at NTC96 did not experience such personnel changes during its pre-NTC Janus training. The untimely 1-163 Battalion/TF staff turbulence, however, should not cause us to lose sight of the remarkable achievement of an ARNG battalion performing at the NTC after being pieced together 3 years ago—from units most of which had no mech infantry experience—and then taking a year out for NET.

C. SUPPORT BATTALION

The NTC96 baseline is limited to relatively few performance observations; however, we conclude that the superior performance of the Support Battalion at NTC98 is the result of its extensive use of Janus and CSS ICW (interactive courseware), both of which were unused by the NTC96 Support Battalion. The performance data and SIMITAR usage data, however, do not enable us to apportion credit between the two interventions.

D. CONCLUSIONS

Compared to NTC96 benchmarks, superior performance at NTC98 is attributed to the amount of SIMITAR training—Janus for the Brigade, Armored TF, and Support Battalion and also CSS ICW for the Support Battalion—that differentiates the two SIMITAR test brigades at NTC98 and NTC96. Comparison of performance on SIMITAR trained tasks with performance on other tasks at both NTC98 and NTC96 provides additional evidence of the effectiveness of Janus training for the Brigade, Armored TF, and Mech Infantry TF.

Superior performance of BFV platoons at NTC98 on Maneuver tasks trained by SIMITAR compared to performance on tasks not trained by SIMITAR is attributed to the ARSI (ARPA Reconfigurable Simulator Initiative) and Platoon Janus Maneuver simulation devices.

I. INTRODUCTION

A. PURPOSE

This report describes an evaluation of the effectiveness of SIMITAR (Simulation in Training for Advanced Readiness). As identified in Table I-1, the SIMITAR Program encompasses various simulation devices, computer-based courseware, and training strategies designed to train tasks described in ARTEP (Army Training and Education Program) MTPs (Mission Training Plans). Table I-2 describes the functions of the "interventions." More details of the SIMITAR Program and the interventions can be found in our SIMITAR Assessment Status Report One (Ref. 1). We sorted the SIMITAR interventions by application into four training areas: Gunnery, Maneuver, BSS (Battle Staff Synchronization), and CSS (Combat Service Support). This assessment report on SIMITAR and NTC98 (National Training Center 98) focuses primarily on BSS and CSS. We are also able to make a limited examination of Platoon Maneuver. An assessment of Gunnery and a more complete assessment of Maneuver will be described in a subsequent overall SIMITAR assessment report, which will also include the BSS and CSS performance results reported here.

B. ASSESSMENT BASELINES

Our examination of SIMITAR involves two assessment baselines. The first baseline provides an external benchmark of training performance by other ARNG (Army National Guard) units at the NTC. The second is an internal baseline in which performance on tasks *not* trained by SIMITAR serve as a benchmark for the performance of tasks trained by SIMITAR.

1. External Benchmark

Army OCs (observer-controllers) provided data for this assessment by rating task performance of the 116th Armored Cavalry Brigade at the NTC in July 1998. The NTC98 performance is compared to performance by the 48th Mechanized Infantry Brigade at NTC in June 1996. These eSBs (enhanced Separate Brigades) of the ARNG are SIMITAR test brigades that differ markedly in the amount of SIMITAR training they had before their NTC rotations—the 48th Brigade had little SIMITAR training and the 116th Brigade had a lot.

At the time it went to the NTC, the 48th Brigade had devoted about 1 percent of its available man-hours to SIMITAR-related training—COFT (Conduct of Fire Trainer) for Gunnery, SIMNET for Maneuver, and Janus for BSS. At that same time, SIMITAR training had accounted for over 3 percent of available man-hours in the 116th Brigade (Ref. 1). And in the 2 years since then, SIMITAR simulators, computer courseware, and training strategies had been the core of the 116th Brigade's training program. We estimate that the 116th Brigade had accumulated over 10 times as many pre-NTC98 SIMITAR training hours as the 48th Brigade had in its pre-NTC96 training program. Thus, we use NTC96 performance as the baseline for evaluating NTC98 performance.

2. Internal Benchmark

That some tasks are trained by SIMITAR devices and courseware while other tasks are not gives us a within-unit benchmarking option. We use it to assess SIMITAR at the brigade, task force/battalion, and company levels for performance at both NTC98 and NTC96. And since platoon training performance was not measured at NTC96, we also use the within-unit benchmark to evaluate platoon performance at NTC98.

C. INDEPENDENT VARIABLES

While we are primarily concerned with one variable, SIMITAR training hours, there are other variables that can affect task performance in the field. One group of variables includes intelligence, education, physical fitness, and experience of brigade personnel. We expect no significant differences between large groups (about 5,000) recruited and promoted to common standards. The 48th Brigade, however, may have an experience advantage, which we did not attempt to quantify: many officers and NCOs went through 8 weeks of post-mobilization training at the NTC after the Reserve call-up during the Gulf War period. The 116th Brigade had no comparable experience. On the other hand, the 116th Brigade benefited in some unmeasured way from lengthy lessons-learned briefings given by a large 48th Brigade contingent to 116th Brigade personnel a few months following NTC96.

Another group of variables includes leadership, unit stability, and cohesion. These factors are difficult to measure, and we did not attempt to measure them. The same observation applies to diligence in each brigade in training both before and during the NTC. We have no a priori reasons to suspect that the brigades differed in these factors.

Another set of variables that could have affected performance was the test conditions. Although the training exercises of interest occurred 2 years apart, both were conducted in hot summer months over the same terrain. Performance of the NTC's opposing force (OPFOR) can also affect performance of units being trained. There is no known difference in OPFOR performance at NTC96 and NTC98.

The competence and/or bias of the OCs could potentially affect recorded performance, if not actual performance. Neither incompetence nor bias was discernible in either training exercise. Standards for task performance and OC selection criteria were unchanged between NTC96 and NTC98. Although all task performance grading was done by OCs at NTC98, OCs graded performance of only the Mechanized Infantry Task Force at NTC96. SIMITAR project personnel—all with substantial Active Army experience in combat arms—graded NTC96 task performance by the Brigade HQ and the Armored Task Force; these

SIMITAR project observers accompanied OCs in the field day and night throughout the entire training exercise.

In summary, comparison of NTC98 and NTC96 performances is not a controlled experiment; however, none of the foregoing unmeasured variables appears to undermine the credibility of this performance comparison. Thus, we believe that observed differences between performances at NTC98 and NTC96 can be attributed to differences in the single independent variable that clearly differentiates the brigades—the amount of SIMITAR training.

D. ORGANIZATION OF THIS REPORT

Chapter II describes our methodology for measuring SIMITAR effectiveness. Our analysis in Chapter III focuses first on the combat elements of the 116th Brigade—principally the HQ staff of the Brigade itself, the Armored TF, and the Mech Infantry TF—and then on the Brigade's combat service support element, the Support Battalion. Chapter IV summarizes our report.

Table I-1. SIMITAR Interventions

AFIST	Abrams Full-Crew Interactive Simulator
ARSI	Advanced Research Projects Agency Reconfigurable Simulator Initiative
ATAFS	Automated Training Analysis and Feedback System
BFVS Gunner Course	Bradley Fighting Vehicle System Gunner's Course
BSTS	Battle Staff Training System
COFT	Conduct of Fire Trainer
Compressed Gunnery	Time-Compressed, Technology-Based Tank Gunnery Training Strategy
CSS ICW	Combat Service Support Interactive Courseware
D–FIRST	Deployable Force-on-Force Instrumented Range System
DSTATS	Digital Systems Test and Training Simulator
EST	Engagement Skills Trainer
GFIST II	Guard Unit Armory Device Full-Crew Interactive Simulator, Field Artillery
Janus (Bde and Bn)	Battle Staff Trainer
Mobile SIMNET	Mobile Simulation Networking
PENCIL	Pen-Based Electronic Network for Command Information Linking
Pile-On Training	Multi-Echelon Training in Several Simulators in a Single Drill Weekend
RCVTP	Reserve Component Virtual Training Program
S–2 Trainer	Intelligence Staff Officer Trainer
SIMBART	Simulation-Based Mounted Brigade Training Program
SIMUTA	Simulation-Based Multiechelon Training Program for Armor Units
SLT	Staff Linkage Trainer
Tank Commander Course	Abrams Tank Commander's Course
Tank Gunner Course	Abrams Tank Gunner's Course
Triage	Voice-Operated Medical Triage Trainer
VMAT	Virtual Reality Maintenance Trainer

Table I-2. SIMITAR Interventions: Functional Descriptions

AFIST	Improvements to an existing low-cost, transportable multimedia tank gunnery simulator that is attached to a stationary tank located in a training area.
ARSI	Reconfigurable simulation platform capable of emulating Abrams and Bradley armored vehicles and High Mobility Multipurpose Wheeled Vehicles (HMMWVs).
ATAFS	Computer-based aid for delivering after-action reviews for SIMNET and ARSI exercises.
BFVS Gunner Course	Complete course of instruction for Bradley gunners. The course is designed to last no more than 6 days.
BSTS	Multimedia courseware for training individual battalion and brigade battle staff officers.
COFT	Existing gunnery training device for Abrams tanks and Bradley Fighting Vehicles.
Compressed Gunnery	Training strategy prescribing the appropriate sequence and mixture of device-based and live-fire gunnery training over an ARNG training year.
CSS ICW	Multimedia-based interactive courseware for training individuals in medical, supply, maintenance, and transport companies and for training support battalion staff members.
D-FIRST	A transportable instrumented range system, based on existing Global Positioning System (GPS), that allows tracking and real-time casualty assessment of up to 78 vehicles in a 20×30 km area.
DSTATS	Tactical device communications simulator that provides the user the ability to interact with field artillery systems [e.g., Multiple Launch Rocket System (MLRS), Joint Surveillance Target Attack Radar System (JSTARS)] and receive/transmit messages according to standard formats.
EST	Computer-based simulation for providing dismounted squads training on coordination and firepower distribution; it also provides limited training and feedback on squad marksmanship.
GFIST II	Device for training Fire Support Specialists (MOS 13F) and other personnel to call for artillery fire.

Table I-2. SIMITAR Interventions: Functional Descriptions (Continued)

Janus (Bde and Bn)	Enhancements to the Janus system significantly reduce its cost, add CS and CSS functions, and provide distributed interactive simulation (DIS) capabilities.
Mobile SIMNET	Existing mobile simulator for training company and platoon maneuver. Enhancements allow it to be used on long-haul network for executing battalion and brigade exercises.
PENCIL	Laptop computers designed to facilitate development and communication of tactical planning products.
Pile-On Training	Training strategy for maximizing the use of gunnery and tactical simulators during an IDT (inactive duty training) weekend drill.
RCVTP	Program for implementing SIMUTA and SIMBART materials that are located at the Fort Knox Mounted Warfare Simulation Training.
S–2 Trainer	Theory-based courseware for training the intelligence officer (S2) in battalions and brigades.
SIMBART	Structured SIMUTA-like program for training Reserve Component (RC) armor brigades.
SIMUTA	Simulation-based structured program for training RC armor battalions, companies, and platoons on Janus and SIMNET.
SLT	Multimedia computer-based program for training staff dyads and triads in the support battalions that are organic to heavy brigades.
Tank Commander Course	Complete course of instruction for Abrams tank commanders designed to last up to 10 days.
Tank Gunner Course	Complete course of instruction for Abrams tank gunners designed to last up to 6 days.
Triage	Multimedia simulation designed to train medics in combat triage. It uses speaker-independent speech recognition, which allows the user to talk with casualties and obtain information.
VMAT	PC-based 3D virtual environment, allowing the student to enter a tank or a BFV and diagnose electrical faults using a virtual version of the Army's simplified test equipment.

II. ASSESSMENT FRAMEWORK

A. TASKS

Army operations in battle involve seven major functions, called BOSs (Battlefield Operating Systems): Intelligence, Maneuver, Fire Support, Air Defense, Command and Control (C2), Mobility and Survivability, and CSS. Because the BOSs are so broad, the Army Research Institute (ARI) developed 39 CCFs (Critical Combat Functions)—a higher resolution set of functions—for assessing training performance within the BOS framework. SIMITAR is a task-training program—SIMITAR interventions do not train collective functions at the BOS or CCF levels. We therefore want to measure performance of tasks (from ARTEP MTPs) included in the BOS and CCF functions.

We directed our evaluation at tasks under the 14 CCFs in which ARNG units were most deficient during field training performance in the 1990–91 Gulf War period; these tasks were the focus for the development of SIMITAR simulation devices and courses.[2] Table II-1 identifies the 14 CCFs and shows the distribution of 518 individual, small unit collective, and battle staff tasks that are included under the BOSs and CCFs.

Most of the 518 tasks are brigade- and battalion-level tasks. The 518 task list includes 403 brigade- and battalion-level tasks, 68 company-level tasks, and 47 platoon-level tasks, most of which (45) apply to Scout and Mortar platoons. In addition to tasks derived from the CCFs, our NTC98 task list also includes

---

2 SIMITAR Program personnel selected priority CCFs by (1) reviewing various Persian Gulf War after action reports; (2) analyzing research reports; (3) interviewing personnel involved in mobilization and training of Roundout brigades; and (4) holding discussions with other relevant groups (Ref. 2).

56 tank platoon tasks and 61 BFV platoon tasks identified in the *Heavy Enhanced Brigade Training Strategy* (Ref. 3).

B. TASKS AND SIMITAR INTERVENTIONS

SIMITAR project personnel identified tasks that were to be trained by the various SIMITAR devices and courses. Our computer file matrix of tasks and interventions is illustrated by Table II-2, where an "X" means a task is trainable by the indicated intervention. (We imply no difference between "SIMITAR trainable" amd "SIMITAR trained" because SIMITAR is used to train most tasks it is capable of training.)

Table II-3 summarizes by BOS the 518 brigade-, battalion-, company-, and platoon-level tasks trainable by SIMITAR interventions. Eighty-five percent of the 518 tasks under the seven BOSs are trainable by at least 1 intervention. Janus alone can train 76 percent of all tasks. ARSI, BSTS, CSS ICW, M–SIMNET, SIMBART, and SIMUTA are other interventions capable of training many tasks.

Table II-4 summarizes the tasks trainable for tank and BFV platoons by various SIMITAR interventions. ARSI, SIMNET(A), SIMNET(B), and P-Janus are the principal training devices for these platoons. The DFIRST, on which over three-fourths of the tasks can be trained, was not available.

C. RATING TASK PERFORMANCE

Table II-5 illustrates by excerpt our computer file containing OC ratings for tasks observed at NTC98. Although company-level tasks and ratings are included under parent task forces (TFs), they are identified separately for analysis of performance at various echelons. Tank and BFV platoon tasks and ratings are not included under parent TFs because platoon-level performance was not rated at NTC96.

D. ANALYTIC METHODS

We used linear and nonlinear methods to analyze the regression of NTC task performance on the SIMITAR units tested at the NTC; these units are a surrogate for amount of SIMITAR training. In the

linear method, we assigned values of 2, 1, and 0 to ratings of Trained (T), Needs Practice (P), and Untrained (U), respectively. The linear method assumes the dependent variable, task performance, is measured on a linear scale; the linear scoring method involves a linear regression analysis to develop a least square line for predicting NTC performance. The OC rating, however, is a discrete variable representing only ordinal relationships; this simple linear scale may not reflect the actual mathematical relationship among the ratings. For example, T + P may not equal 2P.

The use of linear regression is appropriate for continuous variables but can lead to erroneous conclusions about statistical significance when the dependent variable is discrete, such as our case of OC rating (T, P, U) of task performance. Thus, we also use a nonlinear analysis method in which OC ratings are ordinal, i.e., T > P > U, which we established by assigning, as in the linear method, 2, 1, and 0 for T, P, and U, respectively, to start the iterative regression process. In this method, the observed *discrete* dependent variable, OC rating (R), is replaced by an unobserved, *continuous* dependent variable, OC judgment (J). Computer-based regression programs (Ordered Probit and Ordered Logit)[3] substitute the probability of occurrence of a rating for the discrete rating itself and search for a cutoff point (C) between P and T, and for an equation that best predicts OC ratings. The conversion of variables in the regression analysis is shown below.

R	J
U	$-\infty < J \leq 0$
P	$0 < J \leq C$
T	$C < J < \infty$

3 The computer-based regression programs, which are limited dependent variable models, are contained in Ref. 4. Application and discussion of Ordered Probit and Ordered Logit are contained in Refs. 5 and 6.

Table II-1. BOSs, CCFs, and Tasks

BOS	CCF		No. of Tasks	
Intelligence	2	Collect Information	5	
	3	Process Information	3	
	4	Disseminate Information	4	
		Subtotal		12
Maneuver	5	Conduct Tactical Movement	103	
	6	Engage Enemy with Direct Fire and Maneuver	69	
		Subtotal		172
Fire Support	7	Employ Mortars	57	
	15	Coordinate, Synchronize, and Integrate Fire Support	33	
		Subtotal		90
Air Defense	16	Take Active Air Defense Measures	33	
		Subtotal		33
C2	18	Plan for Combat Operations	13	
	19	Direct and Lead Unit During Preparation	5	
	20	Direct and Lead Unit in Execution	6	
		Subtotal		24
Mobility and Survivability	21	Overcome Obstacles	88	
	27	Provide Decontamination	61	
		Subtotal		149
CSS	29	Conduct Supply Operations	38	
		Subtotal		38
7 BOSs	14 CCFs		518 Tasks	

Table II-2. Example Portion of Task vs. Intervention Matrix[a]

BOS	CCF	Task	SIMITAR Interventions				
			AFIST	ARSI	BSTS	Triage	VMAT
Intel	2	1. Analyze TF Order and R&S Plan			X		
Intel	2	2. Prepare for Intelligence Collection			X		
		(Other Tasks and Interventions)					
CSS	29	36. Conduct Immediate/Emergency Resupply					
CSS	29	37. Consolidate and Resupply					
CSS	29	38. Prepare to Continue or Change Mission					

[a] X indicates an intervention can train a task; blank cells indicate no applicability.

Table II-3. Number and Percent of Tasks Trainable by SIMITAR Interventions

Intervention	Battlefield Operating Systems							
	Intelligence	Maneuver	Fire Support	Air Defense	C2	Mobility/ Survivability	CSS	Totals
	Numbers of Tasks							
	12	172	90	33	24	149	38	518
AFIST	2 (17%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0%)
ARSI	5 (42%)	113 (66%)	0 (0%)	0 (0%)	0 (0%)	21 (14%)	0 (0%)	139 (27%)
BSTS	5 (42%)	20 (12%)	15 (17%)	13 (39%)	4 (17%)	32 (21%)	6 (16%)	95 (18%)
COFT	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
CSS ICW	4 (33%)	40 (23%)	0 (0%)	0 (0%)	8 (33%)	0 (0%)	21 (55%)	73 (14%)
DSTATS	1 (8%)	2 (1%)	8 (9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	11 (2%)
EST	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
GFIST II[a]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
Janus	11 (92%)	147 (85%)	73 (81%)	25 (76%)	18 (75%)	90 (60%)	29 (76%)	393 (76%)
M-SIMNET	5 (42%)	109 (63%)	0 (0%)	0 (0%)	0 (0%)	19 (13%)	0 (0%)	133 (26%)
SIMBART	2 (17%)	49 (28%)	46 (51%)	17 (52%)	15 (63%)	30 (20%)	14 (37%)	173 (33%)
SIMUTA	2 (17%)	67 (39%)	53 (59%)	17 (52%)	15 (63%)	32 (21%)	16 (42%)	202 (39%)
SLT	0 (0%)	2 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0%)
TRIAGE	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
VMAT	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Any Intervention[b]	11 (92%)	167 (97%)	75 (83%)	28 (85%)	21 (88%)	105 (70%)	34 (89%)	441 (85%)

[a] Does not train brigade- and battalion-level tasks in Fire Support; trains only one task (Call for Fire) at lower echelons.

[b] Last row indicates number and percent of tasks that are trainable by one or more interventions.

Table II-4. Number and Percent of Tank and BFV Platoon Tasks Trainable by SIMITAR Interventions

Intervention	Tank Platoon	BFV Platoon	Totals
	Numbers of Tasks		
	56	61	117
AFIST	6 (11%)	0 (0%)	6 (5%)
ARSI	31 (55%)	47 (77%)	78 (67%)
ATAFS	5 (9%)	0 (0%)	5 (4%)
BG Course	0 (0%)	17 (28%)	17 (15%)
COFT	2 (4%)	11 (18%)	13 (11%)
CSS ICW	0 (0%)	0 (0%)	0 (0%)
DFIRST	45 (80%)	45 (74%)	90 (77%)
DSTATS	0 (0%)	0 (0%)	0 (0%)
EST	0 (0%)	15 (25%)	15 (13%)
FIST-B[a]	0 (0%)	38 (62%)	38 (32%)
GF II	0 (0%)	2 (3%)	2 (2%)
P-Janus[b]	43 (77%)	44 (72%)	87 (74%)
SIMNET (A)[c]	19 (34%)	0 (0%)	19 (16%)
SIMNET (B)[d]	0 (0%)	22 (36%)	22 (19%)
SLT	0 (0%)	0 (0%)	0 (0%)
TC COURSE	0 (0%)	0 (0%)	0 (0%)
TG COURSE	0 (0%)	0 (0%)	0 (0%)
TRIAGE	1 (2%)	0 (0%)	1 (1%)
VMAT	2 (4%)	1 (2%)	3 (3%)
Any Intervention[e]	53 (95%)	59 (97%)	112 (96%)

[a] FIST-Bradley

[b] Platoon-Janus

[c] SIMNET (Abrams)

[d] SIMNET (Bradley)

[e] Last row indicates number and percent of tasks that are trainable by one or more interventions.

Table II-5. Example Portion of Task Performance Matrix

BOS	CCF	Task	Score[a]				
			Brigade	Armored TF	Mech Inf TF	Company	Platoon
Intel	2	1. Analyze TF Order and R&S Plan					
Intel	2	2. Prepare for Intelligence Collection		1	2		
CSS	29	517. Consolidate and Resupply		1	0		
CSS	29	518. Prepare to Continue or Change Mission		1	1		
		634. Destroy or Abandon M2 BFV					
		635. Perform Before-, During-, and After-Combat Operation Checks					1

[a] Score: Trained (T) = 2; Needs Practice (P) = 1; Untrained (U) = 0. Blank cells indicate tasks were not applicable or their performance was not observed.

III. ANALYSIS AND RESULTS

In our analysis of NTC98, we examine the performance of brigade, TF/battalion, company, and platoon *combat* elements. We use both external and internal baselines to benchmark task performance. After examining performance by these echelons, we focus on brigade and TF performance by BOS. Platoon-level performance at NTC98 is also compared to performance a year earlier at Annual Training in 1997 (AT97). Then we examine the performance of a *combat support* element, the Support Battalion.

A. COMBAT ELEMENTS

Wherever both linear and nonlinear regression methods are used for performance analyses in this chapter, more details—degrees of freedom for the linear regression and test statistics and probabilities of significance for both linear and nonlinear regressions—are shown in Appendix A.

1. Performance by Echelon

Table III-1 indicates Brigade HQ staff performance at NTC98 was much better than NTC96 performance as measured by both linear and nonlinear regression analyses. The results are about the same whether we measure performance on all tasks or on the 85 percent that are trainable by SIMITAR.

Table III-2 shows that the Armored TF at NTC98 performed much better than its counterpart at NTC96 whether we consider all tasks or the reduced set of tasks trainable by SIMITAR. The results are different for the Mech Infantry TFs, where NTC96 performance on SIMITAR trainable tasks was significantly better than NTC98 performance.

Table III-3 shows performance scores on tasks trained by SIMITAR and tasks not trained by SIMITAR for the Brigades and Armored and Mech Inf (Mechanized Infantry) TFs at both NTC98 and NTC96. The table shows average scores and the results of *t* tests for independent samples. We performed these statistical analyses to determine whether the differences in average performance on SIMITAR-trained versus not-SIMITAR-trained tasks were significant (i.e., not due to chance). We tested the hypothesis (two-tailed test) that there is no difference between the average performance scores using the following formula:

$$t = (M_t - M_n)/s_{\text{diff}}$$

where M_t = mean or average for SIMITAR trainable tasks,

M_n = mean or average for SIMITAR nontrainable tasks, and

s_{diff} = estimated standard error of the difference between means.

According to this formula, positive *t* values indicate better performance on SIMITAR trainable tasks. Note that the *t* test statistic considers not only the absolute difference between means of the two sets of task but also the variability and sample size of those two sets. We then determined the probability (*p*) of obtaining a *t* value as large or larger than the calculated value by chance alone. For a given sample size, larger values of *t* (bigger differences) are associated with smaller *p* values (probability of results because of chance). According to statistical conventions, a *t* value is "significant" if the probability of obtaining the result by chance is less than .05 (5 in 100). The *t* test applies to the linear regression method; however, the logic, if not the mathematics, of the *t* test applies to the nonlinear method as well.

Table III-3 shows that the Brigade and the Armored TF at NTC98 and the Mech Inf TF at NTC96 performed significantly better on SIMITAR-trained tasks than on tasks not trained by SIMITAR. At NTC96 the Brigade performance was better, but not significantly, on SIMITAR-trained tasks. The average scores on SIMITAR-trained and not-trained tasks were about equal for the Armored TF at NTC96

(0.63 and 0.66) and the Mech Inf TF at NTC98 (0.63 and 0.65). A discussion of the validity of this within-unit analysis is repeated from our earlier SIMITAR Assessment report (Ref. 1) in Appendix B.

Table III-4 shows linear-regression-derived performance by echelon at NTC98 and NTC96. Performance in both rotations was best at the platoon level and progressively declined at higher echelons. Mean scores were better at all echelons at NTC98. Battalion-level tasks in the table are all Armored TF and Mech Infantry TF tasks minus 68 company-level tasks and 47 platoon-level tasks. Of the latter, 45 of the 47 tasks apply to Scout platoons and Mortar platoons.

In Table III-5, we see that NTC98 performance on SIMITAR-trainable tasks at brigade, battalion, and platoon levels was significantly better than performance at NTC96. For company-level performance, one nonlinear regression gave a result that met our .05, or 5 percent, level of significance test (i.e., the difference in ratings is significant if the probability is less than .05 that the result would occur by chance), while the other nonlinear regression result was significant at the .06 level. The linear regression gave a company-level result that was not significant at the .05 level. However, Table A-4 in Appendix A shows that the probability is less than .08 that the difference in mean scores of 0.83 at NTC98 and 0.72 at NTC96 would occur by chance. Thus, the "no's" in Table III-5's "Statistically Significant" column are actually close to being "yes's."

Performance by tank and BFV platoons was measured at NTC98 but not at NTC96. Table III-6 indicates, by mean OC ratings, that tank platoons performed their tasks a little better than BFV platoons performed theirs. In this comparison, the effects of training and other unit experiences are confounded with differences between tank platoon tasks and BFV platoon tasks.

The activity of both types of platoons involved Maneuver tasks, some of which were SIMITAR trained—by SIMNET (Abrams) for tank platoons and ARSI and P-Janus[4] for BFV platoons—while others

---

4 Platoon-Janus.

were not. Using our within-unit baseline for assessing SIMITAR, we show in Table III-7 that tank platoons scored slightly better (1.12 vs. 1.06), but not significantly better statistically, on tasks not trained by SIMNET (Abrams) than on tasks trained by SIMNET (Abrams). The BFV platoons scored significantly better (0.95 vs. 0.70) on tasks trained by ARSI and P-Janus than on tasks not trained by SIMITAR.

Table III-8 also compares performance within the 116th Brigade. Both linear and nonlinear regression analyses indicate no significant difference in tank platoon and BFV platoon performance at NTC98 and at AT97. These results support the proposition that both types of platoons sustained their AT97 performance in the tougher NTC98 environment.

2. Performance by BOS

The remaining results concern brigade and task force performance on SIMITAR trainable tasks by BOS. Because of the possibility of an erroneous conclusion when using the linear regression for the performance rating variable, we use the nonlinear regression models as our primary test of statistical significance. In cases where the nonlinear model cannot determine statistical significance—due to zero variance (all OC ratings are the same) in either group of data being considered—we use the linear model to determine statistical significance.

Table III-9 shows that Brigade performance in four BOSs—Intelligence, Air Defense, Mobility and Survivability, and CSS—was significantly better at NTC98 than at NTC96. In Table III-10, we see that Armored Task Force performance in four BOSs—Maneuver, Fire Support, C2, and CSS—was significantly better at NTC98 than at NTC96.

Table III-11 shows Mech Infantry TF performance at NTC98 is significantly better than NTC96 performance in only the Fire Support and Air Defense BOSs, unlike Brigade and Armored TF performance. Mech Infantry TF performance at NTC96 is significantly better than NTC98 performance in four BOSs—Maneuver, C2, Mobility and Survivability, and CSS.

3. One Intervention Used

Before attempting to explain the inferior NTC98 SIMITAR performance of the Mech Infantry TF in comparison to performances by the Brigade and Armored TF, we need to identify those SIMITAR interventions that were *actually used* by the 116th Brigade to train tasks from the set of interventions that are *capable* of training tasks. This identification process was applied to Brigade and both TFs for all seven BOSs. As an illustration, of the seven interventions capable of training tasks under the Maneuver BOS (see Table II-3) for the Armored TF only Janus was used (see Table III-12). When we examine all BOSs and all SIMITAR interventions, only Janus, which can train numerous tasks under each BOS, was actually used in 116th Brigade training at the brigade and task force levels prior to NTC98.

4. Brigade HQ and Armored Task Force

The 116th Brigade's overall 10-for-1 advantage in SIMITAR training hours includes Gunnery and Maneuver simulation hours in addition to those related to BSS and CSS. Janus accounted for 3 percent of training hours at the end of April 1996. Since then the 116th Brigade has done two task force/battalion Janus exercises each quarter and a brigade exercise each quarter. Table III-13 shows a more accurate comparison for BSS alone: the number of brigade- and TF-level Janus exercises conducted in the 2 to 3 years prior to the 116th Brigade and the 48th Brigade going to the NTC. A task force/battalion Janus exercise involves about 40 personnel over a weekend drill period (about 16 hours). A brigade Janus exercise involves about 150 personnel (when expenditure and replenishment of fuel and ammunition is emulated) to 225 personnel (when maintenance and medical evacuation functions are also emulated) over a drill weekend. Table III-13 indicates that, whereas the Brigade and Armored TF of the 116th Brigade enjoyed substantial advantages over their counterparts in the 48th Brigade, the advantage for the Mech Infantry TF was notably smaller (8 to 4).

Table III-14 shows that both the Brigade staff and the Armored TF performed significantly better at NTC98 than their counterparts at NTC96. The last column identifies Janus as the SIMITAR intervention that we infer was responsible across all BOSs for the superior training performance.

Training in CSS for the staffs of the Brigade HQ and both TFs included (1) three full CSS Janus exercises in the year and a half prior to NTC98 and (2) limited CSS in other Janus exercises.[5] In our initial draft of this report, which we briefed at the 116th Cavalry Brigade NTC After-Action Review at Biloxi, Mississippi, in late September 1998, we also credited CSS ICW for superior performance in the Maneuver, C2, and CSS BOSs. Review of the training history for the 116th Brigade indicates that credit is not deserved. The CSS ICW training effort was directed only at 116th Brigade's Support Battalion (its Supply and Transportation Company, Maintenance Company, and Medical Company) and not the Support, Maintenance, and Medical Platoons of the maneuver battalions' HQ companies. The effectiveness of the CSS training effort is covered later in this chapter when we compare NTC performance of the Support Battalions of the two SIMITAR test brigades.

5. Mechanized Infantry Task Force

The Mech Infantry TF at NTC98 had an 8-to-4 advantage over its NTC96 counterpart in pre-NTC Janus exercises. However, lack of continuity in the HQ staff of the Mech Infantry TF at NTC98 appears to have nullified its nominal Janus advantage. Numerous changes in 1-163 Battalion staff (which was the Mech Infantry TF staff) were made between AT97 in June 97 and NTC98 in July 98: the S1 and S2 were changed in August 97, the S4 in February 98, the Assistant S3 in March 98, and the Fire Support Officer

---

5 Limited CSS Janus exercises involve emulation of expenditure and replenishment of fuel and ammunition. Besides accounting for fuel and ammunition, full CSS Janus exercises involve emulating (1) use of medical personnel and vehicles to evacuate casualties and (2) use of personnel and material resources to assess battle damage and to replace components, requisition parts, and track the maintenance effort to make repairs.

in May 98.[6] Among them, these six staff members missed over two-thirds of the 1-163 Battalion/TF Janus training. The baseline Mech Infantry TF at NTC96 is reported to have had no battle staff changes throughout four Janus exercises over a 2-year period prior to NTC.[7] The performance results in Table III-15 indicate that, without the Janus advantage, the Mech Infantry TF at NTC98 was significantly superior to its NTC96 counterpart in only the Fire Support and Air Defense BOSs; and the Mech Infantry TF at NTC96 was significantly superior in four BOSs—Maneuver, C2, Mobility and Survivability, and CSS.

Another factor that may have exacerbated the consequences of battle staff instability is the recent formation of the 1-163 Mech Infantry Battalion. In FY95, the Montana ARNG formed 1-163 by rolling up several disparate Guard units, most of which had no mechanized infantry experience; none had BFV experience. After new equipment training during FY96, the battalion had less than 2 years—57 training days—to train for NTC98. In comparison, the NTC96 Brigade HQ and TFs were parts of a long-established brigade with many officers and NCOs who had 8 weeks of post-mobilization training at the NTC in the 1990–91 Gulf War period.

6. Combat Element Summary

Disregarding the Mech Infantry TF results as due to non-SIMITAR effects, the results from NTC98 indicate that SIMITAR training was effective. Examination of the capabilities and usage of SIMITAR interventions establishes a connection to only one intervention—Janus. Evidence related to the effectiveness of other interventions must await analysis of performance at AT, where ratings are focused on platoon performance.

---

6 From 1-163 Bn records (Ref. 7).

7 As recollected by an officer who served with the 48th Brigade during the early 1990's and at NTC96 (Ref. 8).

B. SUPPORT BATTALION

1. Pre-NTC Training

The 116th Brigade's Support Battalion enjoyed a major advantage over the 48th Brigade's Support Battalion in pre-NTC training. The training advantage involves two SIMITAR interventions—Janus (again) and CSS ICW. As revealed earlier, the 116th Brigade's Support Battalion participated in three full-CSS Janus exercises in the year and a half prior to NTC98. Full-CSS Janus had not been developed for the pre-NTC training of the 48th Brigade's Support Battalion. The 116th Brigade's Support Battalion advantage also included substantial use of CSS ICW to train its Supply and Transportation, Maintenance, and Medical Companies in the 6 months prior to NTC98. The 48th Brigade's Support Battalion did not use the CSS ICW.

2. Baseline

Lack of observers to assess SIMITAR at NTC96 (we could not use OCs) limited performance measurements to 21 observations of Support Battalion tasks, all related to the C2 BOS. At NTC98, there were 104 performance observations (by OCs) of tasks related to 4 BOSs—Intelligence, Maneuver, C2, and Mobility and Survivability. Since the Support Battalion task list for NTC96 included no CSS BOS tasks, our baseline for measuring Support Battalion performance is limited to 21 tasks in the C2 BOS.

3. Performance

Table III-16 shows that the Support Battalion performance in the C2 BOS was significantly better at NTC98 than at NTC96. Because the NTC98 ratings have zero variance (all were P's), the nonlinear regression method cannot be used. Some details of the linear regression results are shown in Table III-17. Comparative usage of Janus and CSS ICW and comparative NTC performance indicate these SIMITAR interventions are responsible for the superior performance of the NTC98 Support Battalion. We are not able to apportion credit between the two interventions.

Table III-1. Brigade Staff Performance[a]
(See Appendix A, Table A-1, for more details)

Tasks	Number of Observations		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
All	169	155	4	2	66	26	0.70	0.28	Yes	Yes
SIMITAR Trainable	137	124	4	2	69	27	0.74	0.30	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, U = 0.

e Difference is significant if probability < .05 (two-tailed test) in linear and nonlinear regressions that the result would occur by chance.

Table III-2. Task Force Performance[a]
(See Appendix A, Table A-2, for more details)

Task Force	Tasks	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
		NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Armor	All	958	401	8	4	80	59	0.89	0.63	Yes	Yes
	SIMITAR Trainable	882	351	9	5	81	58	0.90	0.63	Yes	Yes
Mech Inf	All	712	342	5	7	57	61	0.63	0.68	No	No
	SIMITAR Trainable	664	290	5	8	58	64	0.63	0.72	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference is significant if probability < .05 (two-tailed test) in linear and nonlinear regressions that the result would occur by chance.

Table III-3. Brigade and Task Force Performance[a] at NTC98 and NTC96 on Tasks Trained and Not Trained by SIMITAR
(See Appendix A, Table A-3, for more details)

Unit	NTC Rotation	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
		Trained	Not Trained	Trained	Not Trained	Trained	Not Trained	Trained	Not Trained	Linear	Non-linear
Brigade Staff	NTC98	137	32	4	0	69	53	0.74	0.53	Yes	No
	NTC96	124	31	2	0	27	19	0.30	0.19	No	No
Armored TF	NTC98	882	76	9	5	81	71	0.90	0.76	Yes	Yes
	NTC96	351	50	5	0	58	66	0.63	0.66	No	No
Mech Inf TF	NTC98	664	48	5	15	58	50	0.63	0.65	No	No
	NTC96	290	52	8	2	64	40	0.72	0.44	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference is significant if probability < .05 (two-tailed test) in linear and nonlinear regressions that the result would occur by chance.

Table III-4. NTC Performance[a] by Echelon

Echelon	Tasks	Number of Observations		P(T)[b]		P(T or P)[c]		M[d]	
		NTC98	NTC96	NTC98	NTC96	NTC98	NTC96	NTC98	NTC96
Brigade	All	169	155	4	2	66	26	0.70	0.28
	SIMITAR Trainable	137	124	4	2	69	27	0.74	0.30
Battalion	All	1065	596	7	5	65	60	0.72	0.66
	SIMITAR Trainable	966	514	7	6	66	60	0.73	0.66
Company	All	497	75	6	8	77	60	0.83	0.68
	SIMITAR Trainable	492	71	6	8	78	63	0.83	0.72
Platoon[e]	All	108	72	19	6	91	65	1.10	0.71
	SIMITAR Trainable	88	56	18	7	89	70	1.07	0.77

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Most observed tasks (45 of 47) were performed by Scout and Mortar platoons.

Table III-5. NTC Performance[a] on SIMITAR Trainable Tasks by Echelon
(See Appendix A, Table A-4, for more details)

Echelon	Number of Observations		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Brigade	137	124	4	2	69	27	0.74	0.30	Yes	Yes
Battalion	966	514	7	6	66	60	0.73	0.66	Yes	Yes
Company	492	71	6	8	78	63	0.83	0.72	No	No/Yes[f]
Platoon[g]	88	56	18	7	89	70	1.07	0.77	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

f The ordered probit test was not significant ($p = .0614$), whereas the ordered logit test was significant ($p = .0303$).

g Most observed tasks (45 of 47) were performed by Scout and Mortar platoons.

Table III-6. Tank and BFV Platoon Performance[a] at NTC98

Type Platoon	Tasks	Number of Observations	P(T)[b]	P(T or P)[c]	M[d]
Tank	All	470	22	84	1.06
	SIMITAR Trainable	419	22	83	1.06
BFV	All	400	13	79	0.92
	SIMITAR Trainable	357	14	81	0.95

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

Table III-7. Tank and BFV Platoon Performance[a] at NTC98 on Tasks Trained and Not Trained by SIMITAR

(See Appendix A, Table A-5, for more details)

Type Platoon	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	Trained	Not Trained	Trained	Not Trained	Trained	Not Trained	Trained	Not Trained	Linear	Non-linear
Tank	419	51	22	22	83	90	1.06	1.12	No	No
BFV	357	43	14	7	81	63	0.95	0.70	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference is significant if probability < .05 (two-tailed test) in linear and nonlinear regressions that the result would occur by chance.

Table III-8. Tank and BFV Platoon Performance[a] on SIMITAR Trainable Tasks at NTC98 and AT97

(See Appendix A, Table A-6, for more details)

Type Platoons	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	AT 97	NTC 98	AT 97	NTC 98	AT 97	NTC 98	AT 97	Linear	Non-linear
Tank	419	371	22	22	83	90	1.06	1.13	No	No
BFV	357	53	14	9	81	70	0.92	0.79	No	No

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

Table III-9. Brigade Performance[a] on SIMITAR Trainable Tasks by Battlefield Operating System (BOS)
(See Appendix A, Table A-7, for more details)

BOS	Number of Observations		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Intelligence	11	11	0	0	100	36	1.00	0.36	Yes	f
Fire Support	27	21	0	5	42	21	0.42	0.26	No	No
Air Defense	27	14	14	7	75	20	0.89	0.27	Yes	Yes
C2	20	21	5	0	43	55	0.48	0.55	No	No
Mob & Surv	22	35	0	3	78	23	0.78	0.26	Yes	Yes
CSS	30	22	6	0	65	20	0.71	0.20	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

f Nonlinear inestimable due to zero variance in one group.

Table III-10. Armored Task Force Performance[a] on SIMITAR Trainable Tasks by Battlefield Operating System (BOS)
(See Appendix A, Table A-8, for more details)

BOS	Number of Observations		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Intelligence	11	11	0	0	82	45	0.82	0.45	No	No
Maneuver	616	121	8	7	81	61	0.89	0.69	Yes	Yes
Fire Support	75	73	0	5	100	56	1.00	0.62	Yes	f
Air Defense	27	4	0	0	81	75	0.81	0.75	No	No
C2	21	21	81	5	100	57	1.81	0.62	Yes	Yes
Mob & Surv	99	92	2	2	57	50	0.59	0.52	No	No
CSS	33	29	24	0	97	83	1.21	0.83	Yes	Yes

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

f Nonlinear inestimable due to zero variance in one group.

Table III-11. Mech Infantry Task Force Performance[a] on SIMITAR Trainable Tasks by Battlefield Operating System (BOS)
(See Appendix A, Table A-9, for more details)

BOS	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Intelligence	11	11	0	0	36	64	0.36	0.64	No	No
Maneuver	475	104	1	13	56	67	0.57	0.80	Yes	Yes
Fire Support	65	46	22	4	85	63	1.06	0.67	Yes	Yes
Air Defense	27	4	44	0	100	0	1.44	0.00	Yes	f
C2	21	19	0	5	0	42	0.00	0.47	Yes	f
Mob & Surv	33	72	3	10	27	54	0.30	0.64	Yes	Yes
CSS	32	34	0	0	72	100	0.72	1.00	Yes	f

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

f Nonlinear inestimable due to zero variance in one group.

Table III-12. SIMITAR Interventions That Were Used by the 116th Brigade's Armored TF for Maneuver Task Training

SIMITAR Interventions that *can* train tasks under Maneuver BOS	Was the Intervention *used* to train Maneuver tasks?
ARSI	Not at task force level
BSTS	Very little
CSS ICW	Only at Support Battalion
Janus	Yes
M–SIMNET	Not at task force level
SIMBART	Embedded in Janus
SIMUTA	Embedded in Janus and M–SIMNET

Table III-13. Number of Janus Exercises in Pre-NTC Training

Unit	NTC98	NTC96
Brigade[a]	12	1
Armored TF[b]	12	3
Mech Infantry TF[b]	8	4

a A brigade Janus exercise involves 150–225 personnel over an IDT weekend; training time is about 16 hours.

b A task force/battalion Janus exercise involves about 40 personnel over an IDT weekend; training time is about 16 hours.

Table III-14. Superior NTC Performance[a] by Brigade and Armored TF and Related SIMITAR Trainers

Battlefield Operating System	Superior Performance[b]		SIMITAR Trainer
	Brigade, NTC 98	Armored TF, NTC 98	
Intelligence	√		Janus
Maneuver		√	Janus
Fire Support		√	Janus
Air Defense	√		Janus
C2		√	Janus
Mobility and Survivability	√		Janus
CSS	√	√	Janus

a Compared to performance at NTC 96.

b Significance tests in linear and nonlinear regressions (two-tailed tests) indicate probability < .05 that the checked results would occur by chance.

Table III-15. Superior NTC Performance[a] by Mechanized Infantry TFs

Battlefield Operating System	Superior Performance	
	Mech Inf TF, NTC 98	Mech Inf TF, NTC 96
Intelligence		
Maneuver		√
Fire Support	√	
Air Defense	√	
C2		√
Mobility and Survivability		√
CSS		√

[a] Compared to performance at other NTC rotation.

Table III-16. Support Battalion Performance[a] on SIMITAR Trainable Tasks by BOS

BOS	Number of Rated Tasks		P(T)[b]		P(T or P)[c]		M[d]		Statistically Significant?[e]	
	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	NTC 98	NTC 96	Linear	Non-linear
Intelligence	9	0	0	---	100	---	1.00	---	---	---
Maneuver	53	0	0	---	100	---	1.00	---	---	---
C2	11	21	0	0	100	57	1.00	0.57	Yes	f
Mob & Surv	31	0	0	---	100	---	1.00	---	---	---

a Performance is rated Trained (T), Needs Practice (P), or Untrained (U).

b Percent of tasks rated T.

c Percent of tasks rated T or P.

d Mean rating where T = 2, P = 1, and U = 0.

e Difference in ratings is significant if probability < .05 in linear and nonlinear regressions (two-tailed tests) that the result would occur by chance.

f Nonlinear inestimable due to zero variance in one group.

Table III-17. Comparison of Support Battalion Performance at NTC98 and NTC96 by Linear Regression Method

Tasks	Linear Regression		
	df [a]	*t* [b]	*p* [c]
SIMITAR Trainable	153	6.540	.000
C2 BOS	123	8.761	.000

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

IV. SUMMARY AND CONCLUSIONS

SIMITAR involves the use of about 20 interventions—simulation devices, computer-based courseware, and training strategies—to train tasks identified in ARTEP MTPs. We use two baselines for assessing SIMITAR effectiveness. The first baseline, which provides an external benchmark, is the performance of other ARNG units that trained at the NTC. An internal (within-unit) baseline involves performance on tasks *not* trained by SIMITAR to benchmark the performance of tasks trained by SIMITAR.

Our evaluation of two SIMITAR test brigades involves the task performance that was graded by Army OCs at the NTC. The brigades differed markedly in the amount of SIMITAR training before their NTC rotations. We used linear and nonlinear methods to regress task performance grades at the NTC on the SIMITAR test brigades, which were surrogates for the very different amounts of SIMITAR training. The linear method provided numerical results by assigning values of 2, 1, and 0 to the T, P, and U grades, respectively, in the Army's 3-point performance rating system. Because the use of linear regression can lead to erroneous conclusions about statistical significance—when OC grades, which are a discrete, ordinal variable, are converted to a simple linear scale—we also used the nonlinear method. The two regression methods gave similar results throughout our analysis.

The performances of four major Brigade elements—Brigade HQ, Armored TF, Mech Infantry TF, and Support Battalion—were measured at NTC98 and compared to baseline performances of their counterparts at NTC96. The three elements—Brigade HQ, Armored TF, and Support Battalion—that had markedly more pre-NTC Janus training (all three elements) and CSS ICW training (the Support Battalion) performed significantly better than the counterpart elements at NTC96 (see Tables III-14 and III-16). The NTC98 Mech Infantry TF, whose nominal advantage in Janus training was nullified by numerous battle

staff changes, performed significantly better in two BOSs and significantly worse in four BOSs than the NTC96 Mech Infantry TF, which, because of the other's staff changes, may have had an advantage in Janus training (see Table III-15).

In every case in which we used within-unit baselines, average performance at NTC98—and NTC96 also—on SIMITAR-trained tasks was about equal to or significantly better than on tasks not trained by SIMITAR (see Tables III-3 and III-7).

Performance by echelon at both NTC98 and NTC96 was better at the platoon level and progressively declined at company-, battalion-, and brigade-level echelons. Performance at NTC98 was much better at all echelons than at NTC96 (see Tables III-4 and III-5).

Compared to NTC96 benchmarks, superior performance at NTC98 is attributed to the amount of SIMITAR training—Janus for the Brigade, Armored TF, and Support Battalion and also CSS ICW for the Support Battalion—that differentiates the two SIMITAR test brigades at NTC98 and NTC96. Comparison of performance on SIMITAR-trained tasks with performance on other tasks at both NTC98 and NTC96 provides additional evidence of the effectiveness of Janus training for the Brigade, Armored TF, and Mech Infantry TF.

A much smaller set of observations (relative to those at Brigade and TF levels) of Platoon Maneuver performance at NTC98 indicates that BFV platoons performed significantly better on Maneuver tasks trained by SIMITAR interventions—ARSI and Platoon Janus—than on tasks not trained by SIMITAR. Tank platoons scored slightly better, but not significantly better statistically, on tasks *not* trained by SIMNET (Abrams) than on tasks trained by SIMNET (Abrams).

In conclusion, the data from NTC98, in almost every instance, indicate that performance was improved with exposure to SIMITAR training devices, courseware, and strategies. Whether we use external or internal performance benchmarks, our evaluations consistently confirmed the effectiveness of SIMITAR interventions.

REFERENCES

1. Orlansky, J., J. Metzko, J. Morrison, and G. Pickell, *Assessment of SIMITAR: Status Report One*, Institute for Defense Analyses, IDA Document D-2069, April 1997.
2. Krug, Colonel R., and M. Shaler, *SIMITAR Assessment Plan*, briefing for G-3, U.S. Army Forces Command, 20 March 1996.
3. *Heavy Enhanced Brigade Training Strategy*, Coordinating Draft, U.S. Army Armor Center and Army National Guard, January 1995.
4. Green, W.H., *LIMDEP (Limited Dependent) Version 7.0 Users Manual*, Chapter 23, "Ordered Probability Models," Econometric Software, Inc., 15 Gloria Place, Plainview, NY 11803, January 1998.
5. McKelvey, R.D., and W. Zavoina, "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," *The Journal of Mathematical Sociology*, Vol. 4, pp. 103–120, 1975.
6. Spector, L.C., and M. Mazzeo, "Probit Analysis and Economic Education," *The Journal of Economic Education*, Spring 1980.
7. Telephone conversations with Major Lee Fretwell, Training Support Battalion, assigned to 1-163 Mech Infantry Battalion, Montana ARNG.
8. Telephone conversations with Major Al Fracker, U.S. Army Infantry School, Fort Benning, GA.

GLOSSARY

AFIST	Abrams Full-Crew Interactive Simulator
ARI	Army Research Institute
ARNG	Army National Guard
ARPA	Advanced Research Projects Agency
ARSI	ARPA Reconfigurable Simulator Initiative
ARTEP	Army Training and Education Program
AT	annual training
ATAFS	Automated Training Analysis and Feedback System
BFV	Bradley Fighting Vehicle
BFVS	Bradley Fighting Vehicle System
BOS	Battlefield Operating System
BSS	Battle Staff Synchronization
BSTS	Battle Staff Training System
C	cut-off point between T and P in nonlinear regression analysis
C2	Command and Control
CCF	critical combat function

COFT	Conduct of Fire Trainer
CSS	Combat Service Support
D–FIRST	Deployable Force-on-Force Instrumented Range System
DIS	distributed interactive simulation
DSTATS	Digital Systems Test and Training Simulator
eSB	enhanced Separate Brigade
EST	Engagement Skills Trainer
GFIST II	Guard Unit Armory Device Full-Crew Interactive Simulator, Field Artillery
GPS	Global Positioning System
HMMWV	High Mobility Multipurpose Wheeled Vehicle
HQ	headquarters
ICW	interactive courseware
IDT	inactive duty training
J	judgment
JSTARS	Joint Surveillance Target Attack
Mech Inf	Mechanized Infantry
MLRS	Multiple Launch Rocket System
Mob	Mobility
MTP	Mission Training Plan
NCO	noncommissioned officer

NTC	National Training Center
OC	observer controller
OPFOR	opposing force
P	Needs Practice or percentage
PENCIL	Pen-Based Electronic Network for Command Information Linking
R	rating
RC	Reserve Component
RCVTP	Reserve Component Virtual Training Program
SIMBART	Simulation-Based Mounted Brigade Program
SIMITAR	Simulation in Training for Advanced Readiness
SIMUTA	Simulation-Based Multiechelon Training Program for Armor Units
SLT	Staff Linkage Trainer
Surv	Survivability
T	Trained
TF	task force
U	Untrained
VMAT	Virtual Reality Maintenance Trainer

APPENDIX A

Comparison of Performance by Linear and Nonlinear Regression Methods

Table A-1. Comparison of Brigade Performance Between NTC98 and NTC96 by Linear and Nonlinear Regression Methods

Tasks	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
All	322	7.378	.000	7.041	.000	7.001	.000
SIMITAR Trainable	259	6.786	.000	6.487	.000	6.474	.000

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

Table A-2. Comparison of Task Force Performance Between NTC98 and NTC96 by Linear and Nonlinear Regression Methods

Task Force	Tasks	Linear Regression			Nonlinear Regression			
					Ordered Probit		Ordered Logit	
		df[a]	*t*[b]	*p*[c]	*z*[b]	*p*[c]	*z*[b]	*p*[c]
Armor	All	1357	8.014	.000	8.038	.000	8.046	.000
	SIMITAR Trainable	1231	7.969	.000	8.074	.000	8.128	.000
Mech Inf	All	1052	–1.413	.158	–1.415	.157	–1.378	.168
	SIMITAR Trainable	952	–2.416	.016	–2.428	.015	–2.334	.020

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC 96 performance are equal and the observed difference is due to chance.

Table A-3. Comparison of Brigade and Task Force Performance at NTC98 and NTC96 Between SIMITAR Trainable and Nontrainable Tasks by Linear and Nonlinear Regression Methods

	NTC Rotation	Linear Regression			Nonlinear Regression			
					Ordered Probit		Ordered Logit	
		df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Brigade	NTC98	167	1.988	.048	1.861	.063	1.828	.068
	NTC96	153	1.065	.289	1.010	.313	0.937	.349
Armor TF	NTC98	956	2.163	.031	2.180	.029	2.199	.028
	NTC96	399	–0.359	.720	–0.265	.791	–0.523	.601
Mech Inf TF	NTC98	710	–0.237	.812	–0.299	.765	0.270	.787
	NTC96	340	3.167	.002	3.123	.002	3.039	.002

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

Table A-4. Comparison of Performance by Echelon Between NTC98 and NTC96 on SIMITAR Trainable Tasks by Linear and Nonlinear Regression Methods

Echelon	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Brigade	259	6.786	.000	6.487	.000	6.474	.000
Battalion	1478	2.252	.024	2.259	.024	2.323	.020
Company	561	1.748	.080	1.870	.062	2.166	.030
Platoon	142	3.172	.002	3.135	.002	3.080	.002

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

Table A-5. Comparison of Performance at NTC98 on Tank and BFV Platoon Tasks Between SIMITAR Trainable and Nontrainable Tasks by Linear and Nonlinear Regression Methods

Type Platoon	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Tank	468	–0.660	.509	–0.599	.549	–0.543	.587
BFV	398	2.729	.007	2.818	.005	2.838	.005

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

Table A-6. Comparison of Tank and BFV Platoon Performance Between NTC98 and AT97 on SIMITAR Trainable Tasks by Linear and Nonlinear Regression Methods

Type Platoon	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Tank	788	1.640	.101	1.626	.104	1.512	.131
BFV	408	−1.864	.063	−1.923	.054	−1.955	.051

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

Table A-7. Comparison of Brigade Performance by BOS Between NTC98 and NTC96 on SIMITAR Trainable Tasks by Linear and Nonlinear Regression Methods

BOS	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Intelligence	20	5.164	.000	d	d	d	d
Fire Support	46	1.264	.213	1.090	.276	1.417	.156
Air Defense	39	3.165	.003	3.381	.001	3.228	.001
C2	39	–0.439	.663	–0.398	.691	–0.619	.536
Mob & Surv	55	5.737	.000	2.699	.007	2.962	.003
CSS	50	4.517	.000	3.859	.000	3.640	.000

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

d Nonlinear solutions are inestimable due to zero variance in one group.

Table A-8. Comparison of Armored Task Force Performance by BOS Between NTC98 and NTC96 on SIMITAR Trainable Tasks by Linear and Nonlinear Regression Methods

BOS	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Intelligence	20	1.826	.083	1.760	.078	1.705	.088
Maneuver	735	3.940	.000	4.227	.000	4.468	.000
Fire Support	146	5.609	.000	[d]	[d]	[d]	[d]
Air Defense	29	0.297	.769	0.300	.764	0.305	.760
C2	40	7.643	.000	4.534	.000	3.848	.000
Mob & Surv	189	0.821	.413	0.810	.418	0.868	.385
CSS	60	3.428	.001	2.667	.008	2.425	.015

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

d Nonlinear solutions are inestimable due to zero variance in one group.

Table A-9. Comparison of Mech Infantry Task Force Performance by BOS Between NTC98 and NTC96 on SIMITAR Trainable Tasks by Linear and Nonlinear Regression Methods

BOS	Linear Regression			Nonlinear Regression			
				Ordered Probit		Ordered Logit	
	df [a]	*t* [b]	*p* [c]	*z* [b]	*p* [c]	*z* [b]	*p* [c]
Intelligence	20	–1.268	.219	–1.277	.202	–1.263	.207
Maneuver	577	–3.916	.000	–4.175	.000	–3.568	.000
Fire Support	109	3.413	.001	3.219	.001	3.097	.002
Air Defense	29	5.623	.000	d	d	d	d
C2	38	–3.553	.001	d	d	d	d
Mob & Surv	103	–2.577	.011	–2.638	.008	–2.045	.041
CSS	64	–3.592	.001	d	d	d	d

a *df* is degrees of freedom, which equals the number of independent observations in the sample.

b *t* is a test statistic in small sample (N < 30) distributions and increasingly approximates the *z* test statistic for normal distributions as sample size increases.

c *p* is the probability in two-tailed test of significance that NTC98 performance and NTC96 performance are equal and the observed difference is due to chance.

d Nonlinear solutions are inestimable due to zero variance in one group.

APPENDIX B

Validity of Analysis of Tasks Trained and Not Trained by SIMITAR

APPENDIX B
Validity of Analysis of Tasks Trained and Not Trained by SIMITAR

There is no a priori reason to believe that any of the nontraining variables are systematically correlated with the differences between SIMITAR trained and SIMITAR nontrained tasks. To that extent, they may be regarded as sources of random variability. Random variability does not invalidate these comparisons; it can, however, reduce the sensitivity of the analysis to detect actual effects in the data. While the effect of random variability increases the probability of making a false negative decision (declaring a difference not significant when the actual difference is actually significant), it reduces the probability of a false positive decision (declaring a difference significant when there is no actual difference). In short, the effect of these sources of variability was to make the analyses more conservative but not invalid.

A second, and perhaps more serious, threat to the validity of these analyses is variables that are correlated or confounded with differences between SIMITAR trained and other tasks. Perhaps the most obvious confound is the possibility that SIMITAR-trained tasks are inherently easier than other tasks. This could explain why task performance of the Mech Inf TF was better on SIMITAR-trained tasks than on tasks not trained by SIMITAR.* However, the results from the Armored TF and the Brigade staff do not support this argument: they show no significant differences between SIMITAR-trained and nontrained

---

* This discussion pertains to analysis of training performance by units of the 48th Mechanized Infantry Brigade at the National Training Center in 1996 (Orlansky, J., J. Metzko, J. Morrison, and G. Pickell, *Assessment of SIMITAR: Status Report One*, Institute for Defense Analyses, IDA Document D-2069, April 1997).

tasks. In other words, the lack of differences in the other two brigade elements argues against this source of confounding and in favor of the conclusion that SIMITAR training improves NTC task performance.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1998	Final—June 1998–October 1998

4. TITLE AND SUBTITLE
Assessment of SIMITAR: Status Report Two (NTC98)

5. FUNDING NUMBERS
DASW01 94 C 0054
A-132

6. AUTHOR(S)
John Metzko, John Morrison

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Document D-2217

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency/ISO
3701 N. Fairfax Drive
Arlington, VA 22210

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 180 words)***

This report describes an evaluation of the effectiveness of SIMITAR (Simulation in Training for Advanced Readiness). The SIMITAR Program encompasses various interventions, i.e., simulation devices, computer-based courseware, and training strategies designed to train tasks described in ARTEP (Army Training and Education Program) MTPs (Mission Training Plans). The SIMITAR interventions are used in four training areas: Gunnery, Maneuver, BSS (Battle Staff Synchronization), and CSS (Combat Service Support). This assessment report on the effectiveness of SIMITAR in the performance of the 116th Armored Cavalry Brigade at the National Training Center in July 1998 focuses primarily on BSS and CSS; it also includes a limited examination of Platoon Maneuver. An assessment of Gunnery and a more complete assessment of Maneuver will be described in a subsequent overall SIMITAR assessment report, which will also include BSS and CSS performance results reported here.

14. SUBJECT TERMS Army National Guard, Combat Service Support, Compressed Gunnery Strategy, Battle Staff Synchronization, Maneuver training, Pile-on training, SIMITAR, Simulation, training, training courseware, training devices, training performance, training procedures, training readiness, training strategy, training technology

15. NUMBER OF PAGES
80

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR